Condition classification for bearing fault based on machine learning using GUI

Chouaib SOUAIDIA
LABGET Laboratory
Department of Electrical Engineering
Faculty of Science and Technology
Larbi Tebessi University
Tebessa, Algeria
chouaib.souaidia@univ-tebessa.dz

Tawfik THELAIDJIA
LABGET Laboratory
Department of Electrical Engineering
Faculty of Science and Technology
Larbi Tebessi University
Tebessa, Algeria
tawfik.thlaidjia@univ-tebessa.dz

Salah CHENIKHER
LABGET Laboratory
Department of Electrical Engineering
Faculty of Science and Technology
Larbi Tebessi University
Tebessa, Algeria
salah.chenikher@univ-tebessa.dz

Abstract – In this paper, a new method based on a graphical user interface is developed for bearing fault diagnosis. The suggested method consists of using statistical parameters for feature extraction. Then the classification task is guaranteed using two classifiers namely: neural networks and random forest. The suggested approach is tested based on the bearing dataset provided by the Case Western Reserve University, Bearing Data Center. To facilitate the exploitation of the proposed approach a GUI has been developed. The obtained results show the effectiveness and the simplicity of the proposed approach for bearing fault diagnosis.

Keywords: *Rolling Bearing, Fault Diagnosis, ANN, Random forest, Feature Extraction, Classification Algorithms.*

I. INTRODUCTION

Rotating machines are necessary and familiar in the industry. Rolling bearing is the crucial component that directly influences the running condition of the industrial process [1]. Statistics show that more than two-thirds of electrical motor failures are due to rolling bearing damages [2]. Besides, the failure of this mechanic component may cause a catastrophic breakdown of the system and as consequence an expensive maintenance cost [3]. To avoid this shortcoming researchers should be focused on bearing fault diagnosis.

In this paper, a pattern recognition technique is used for bearing fault diagnosis. It is based on key steps: feature extraction and condition classification [4].

Diversity of data analysis techniques have been used by researchers, namely time-domain analysis [6–8], frequency domain analysis [9, 10], and time-frequency domain analysis [11]. Among these features, statistical parameters such as Root mean square (RMS), kurtosis, crest factor, average amplitude, minimum value, standard deviation, and Impulse Factor have been successfully used for feature extraction. These statistical parameters are useful due to their sensitivity to any change that occurs on the vibration signal [4].

Researchers have employed many artificial intelligence techniques for bearing fault classification such as fuzzy logic [12], expert system [12], random forest techniques, neural network (NN) [6, 7, 14, 15], neuro-fuzzy system [8], genetic algorithm (GA) [6, 15]and support vector machine (SVM) [7, 13]. neural networks and random forest techniques are important classification methods, due to their efficiency and accuracy.

The aim of this work is not only to propose an approach for fault diagnosis but also to develop a graphical user interface (GUI) based on MATLAB to ensure simple exploitation by the user.

This paper is organized as follows. In section 2, the feature extraction technique is presented. Then, neural networks and random forest techniques are discussed in section 3. The used dataset is given in section 4. In section 5, the obtained results are analyzed and discussed.

II. FEATURE EXTRACTION

Features extraction plays an important role in data analysis and classification. It allows extracting useful information from signals. Extracting effective parameters may increase the reliability of fault diagnosis. In this paper, seven (7) time-domain features were extracted from vibration signals. The considered parameters are the minimum (Min), the impulse factor (IMF), the crest factor (CRF), standard deviation (SD), the root-mean-square (RMS), the peak to a peak value

(PPV), and the kurtosis (Kur). All these features are given by their expressions as follows:

Peak value: [23]

$$x_p = maximum\ (x) \tag{1}$$

minimum value: [22]

$$Min = minimum\ (x) \tag{2}$$

peak to a peak value (PPV): [22]

$$PPV = maximum(x) - minimum\ (x) \tag{3}$$

root mean square (RMS): [22]

$$x_{RMS} = \sqrt{\frac{\sum_{i=1}^{K} |x_i|^2}{K}} \tag{4}$$

standard deviation (SD): [22]

$$\sigma = \sqrt{\frac{\sum_{i=1}^{K} (x(i) - \bar{x})^2}{K-1}} \tag{5}$$

the crest factor (CRF): [23]

$$CRF = \frac{x_p}{x_{RMS}} \tag{6}$$

impulse factor (IMF): [23]

$$IMF = \frac{x_p}{\frac{1}{K}\sum_{i=1}^{K} |x_i|} \tag{7}$$

kurtosis (KUR) : [23]

$$KUR = \sqrt{\frac{\sum_{i=1}^{K} (x(i) - \bar{x})^4}{K * (\sigma^2)}} \tag{8}$$

Where:

x represents the vibration signal in the time domain.

σ represents the Standard Deviation

K the length of a signal

III. MACHINE LEARNING METHODS

A. ARTIFICIAL NEURAL NETWORKS

Artificial neural networks (ANNs) are biologically inspired computational networks. ANN has the capability of establishing the relationship between the input and output patterns. [17]. It can be configured for a specific application, such as pattern recognition or data classification, through a learning process [18].

Among different topologies, the feed-forward network is probably the most popular; the basic multilayer feed-forward network contains one input layer, one output layer, and several hidden layers [19], [20].

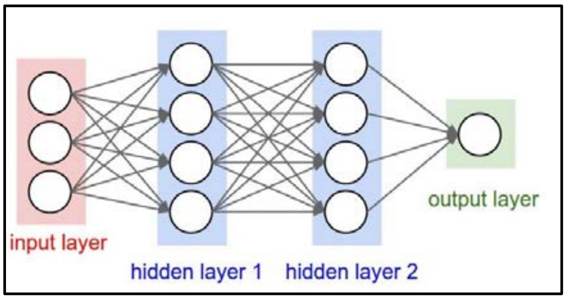

Fig. 1. feed-forward neural network [21]

The sigmoid activation function is used in ANN. It is given by equation 8. The activation function confines the amplitude neuron output to some finite value and introduces non-linearity in the network.

$$Z(x) = \frac{1}{1 + e^{-x}} \tag{9}$$

An ANN model consists of two phases: the learning or training phase and the testing phase. The dataset is divided into two parts; the first one which contains 75% of data is used for ANN training and the rest is employed for testing purposes.

B. RANDOM FOREST

random forest classifier is a tree-based machine learning algorithm that consists of many individual decision trees that operate as an ensemble or group. It uses averaging to improve the predictive or classification accuracy. It can be used both for classification and regression and the more trees it has, the more robust a forest is.

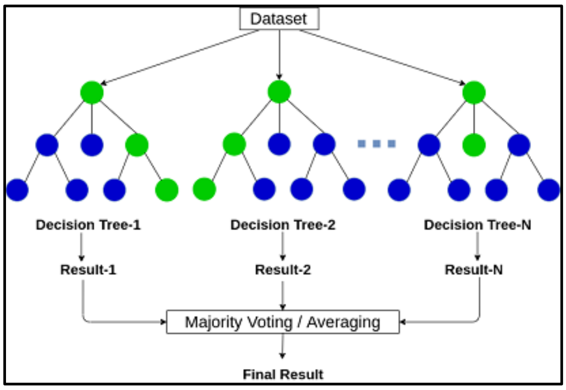

Fig. 2. random-forest Classifier [24]

Random-forest works based on four main steps:

1. Selecting random samples from a given dataset.
2. Creating a decision tree for each sample and get a prediction result from each decision tree.

3. Performing a vote for each predicted result.
4. Finally, selecting the prediction result with the most votes as the final prediction.

IV. DATASET & GUI

A. THE DATASET

The dataset is provided by the Case Western Reserve University Bearing Data Center is used for this study [25].

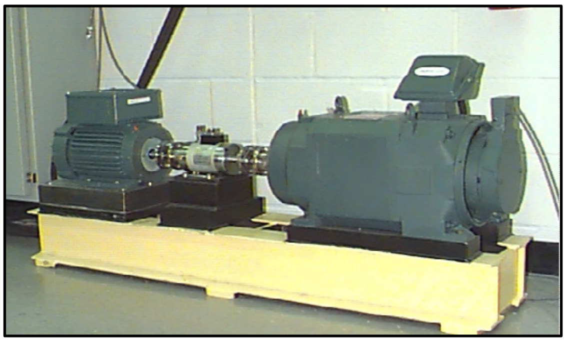

Fig. 3. The test stands (Source: Case Western Reserve University Bearing Data Center website) [25]

As shown in Fig. 3, the test stand consists of a 2 hp motor (left), a torque transducer/encoder (center), a dynamometer (right), and control electronics (not shown).

The following seven bearing conditions are considered in this work:

- Bearing with ball fault of 0.1778 mm (BF1)
- Bearing with ball fault of 0.5334 mm (BF2)
- Bearing with inner race fault of 0.1778 mm (IF1)
- Bearing with inner race fault of 0.5334 mm (IF2)
- Bearing with outer race fault of 0.1778 mm (OF1)
- Bearing with outer race fault of 0.5334 mm (OF2)
- Healthy bearing (HTY).

The dataset used in this study contains 280 samples, each category contains 40 signals.

B. GUI

The developed GUI allows doing the following operations:

1. Choose the dataset folder using the button 'Browse'.
2. Select the desired classifier namely: the feed-forward neural network or the random forest.
3. When the user selects a feed-forward neural network, he can choose the number of hidden layers and the number of epochs as shown in Fig. 7. On the other hand, when the user selects the random forest classifier, he can choose the number of trees as illustrated in Fig. 11.
4. Train the selected classifier once only using the button 'NEW'.
5. Calculate the test using the button 'validate'.
6. The user can check a random signal to which class it belongs by clicking on the check button, it also allows to visualize the values of statistical parameters of the selected signal.

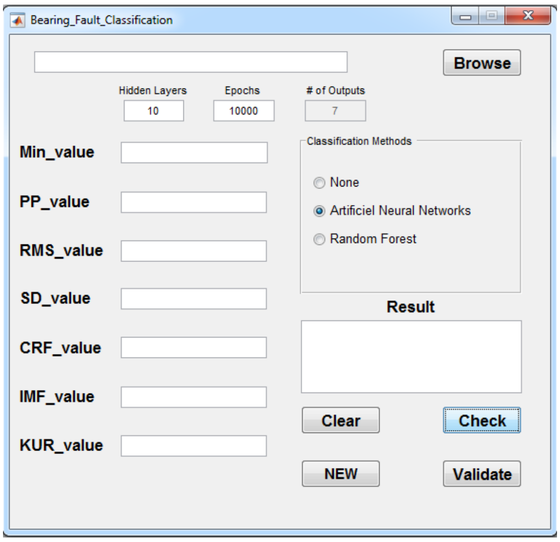

Fig. 4. The developed graphical user interface for classification

This GUI uses the extracted feature as a dataset, in our study the rows are the 280 bearing conditions and columns are the 7 extracted features (statical parameters).

TABLE I. GUI DATASET FORMAT

	Min	PPV	RMS	SD	CRF	IMF	KUR
BF1	-0.520	0.977	0.136	0.136	7.139	8.884	2.877
BF2	-0.411	0.995	0.118	0.117	8.402	10.81	3.580
IF1	-1.172	2.738	0.290	0.289	9.433	13.15	5.588
IF2	-2.347	5.287	0.516	0.515	10.24	15.02	7.185
OF1	-3.212	6.705	0.678	0.677	9.882	16.31	7.621
OF2	-5.427	11.49	0.627	0.627	18.31	36.86	22.40
HTY	-0.272	0.495	0.076	0.075	6.507	8.123	2.891

V. OBTAINED RESULTS

A. *ANN RESULTS:*

After feature extraction, the first classifier based on a feed-forward neural network is employed to classify samples. Fig. 5 illustrates the suggested method for fault diagnosis based on a feed-forward neural network.

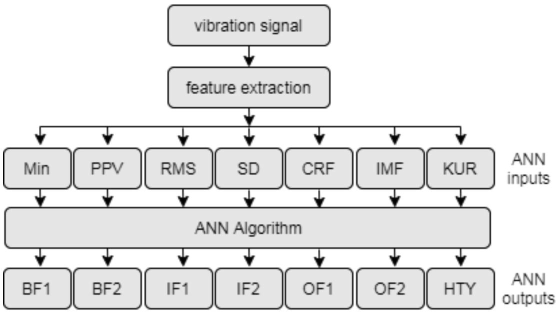

Fig. 5. The suggested method for fault diagnosis is based on a feed-forward neural network.

The inputs of ANN are the seven statistical parameters [Min; PPV; RMS; SD; CRF; IMF; KUR and the outputs are the bearing states: [BF1; BF2; IF1; IF2; OF1; OF2; HTY] as shown in Table II.

TABLE II. DESIRED OUTPUTS OF ANN

Bearing state	Desired Outputs						
ball fault of 0.1778 mm (BF1)	1	0	0	0	0	0	0
ball fault of 0.5334 mm (BF2)	0	1	0	0	0	0	0
inner race fault of 0.1778 mm (IF1)	0	0	1	0	0	0	0
inner race fault of 0.5334 mm (IF2)	0	0	0	1	0	0	0
outer race fault of 0.1778 mm (OF1)	0	0	0	0	1	0	0
outer race fault of 0.5334 mm (OF2)	0	0	0	0	0	1	0
Healthy bearing (HTY).	0	0	0	0	0	0	1

Feed-forward neural network suffers from the main shortcoming which is the correct choice of its architecture. After several experiments, the architecture of the considered feed-forward neural networks is presented in Table III.

TABLE III. THE ARCHITECTURE OF THE CONSIDERED FEED-FORWARD NEURAL NETWORK

Number of hidden layers	10
Maximum number of epochs	10000
Mean square error	10^{-10}

The feed-forward neural network training contains 7 Inputs, 10 Hidden Layers, and 7 outputs as presented in Fig. 6.

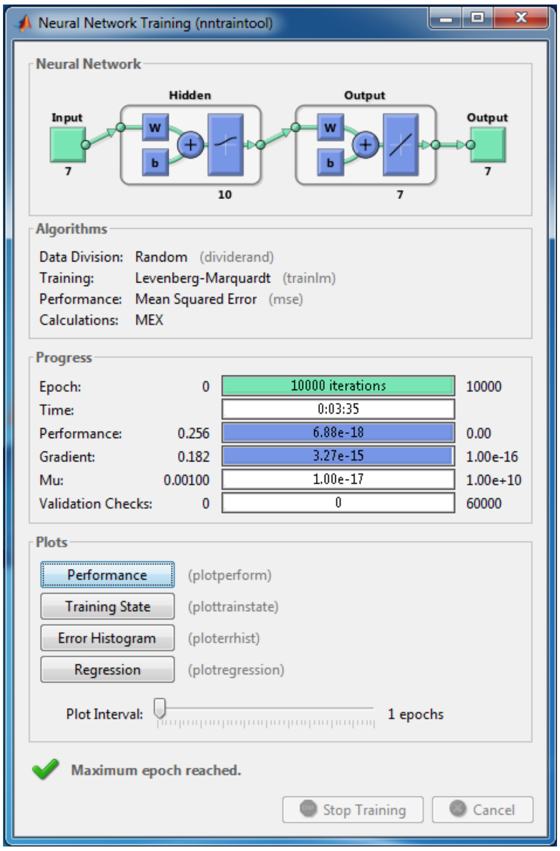

Fig. 6. The feed-forward neural network training

The test accuracy of the optimal feed-forward neural network has reached 100%.

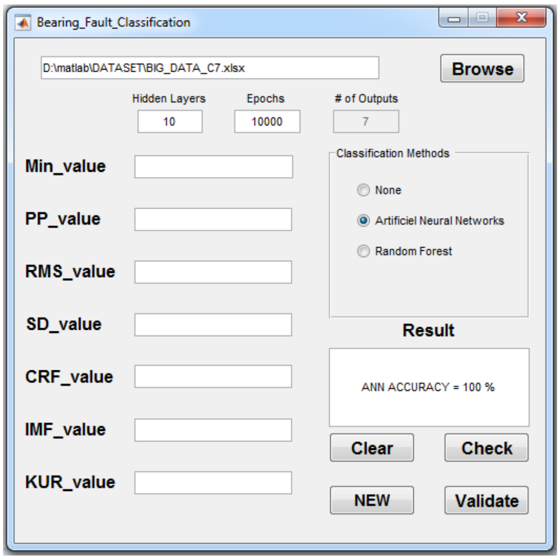

Fig. 7. Accuracy rate based on feed-forward neural network

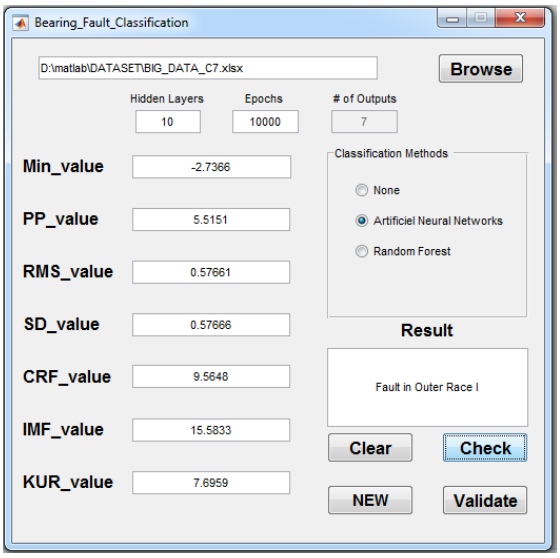

Fig. 8. An example of testing a bearing with Outer Race fault (OF1).

B. RANDOM FOREST RESULTS:

In the case of the random forest classifier, several experiments have been done using a variable number of trees. The obtained results demonstrate that test accuracy is heavily depending on the right choice of the number of trees. Fig. 10 gives the varying error rate according to the number of trees. From Fig. 10 it can be seen that the error rate decrease when the number of trees increases, then it will be established when the number of trees is equal to or greater than 10.

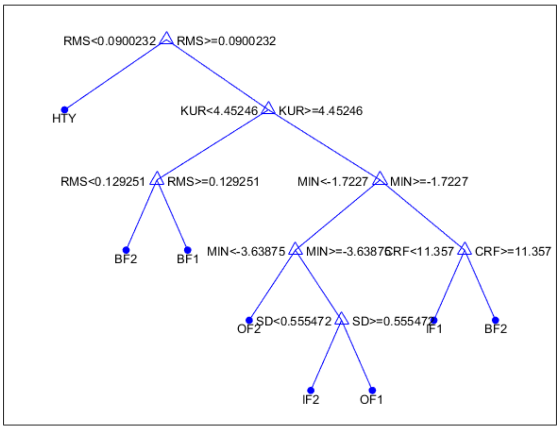

Fig. 9. Classification Tree

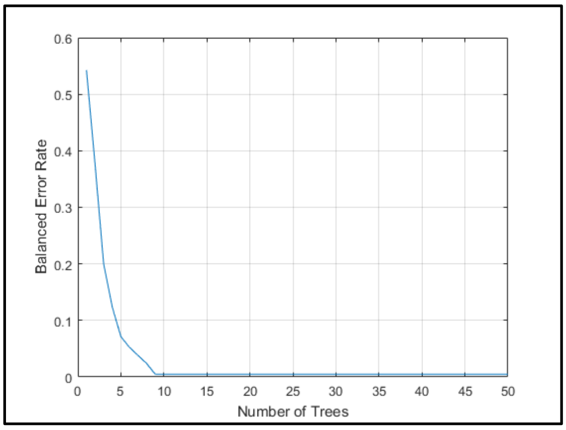

Fig. 10. Varying of the error rate according to the number of trees

The test accuracy obtained using 50 trees is equal to 98.5714%. Fig. 11 shows the developed graphical user interface (GUI) for bearing fault diagnosis.

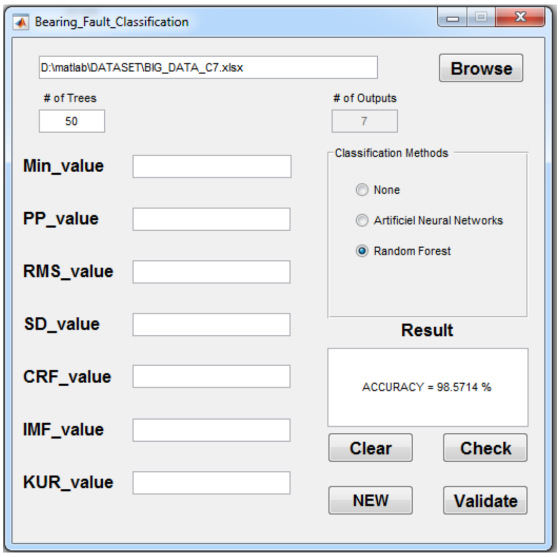

Fig. 11. Accuracy rate based on random forest.

VI. CONCLUSION

In this study condition classification using a graphical user interface has been developed for fault diagnosis. The proposed approach employed statistical parameters for feature extraction. Then, two classifiers namely: feed-forward neural network and random forest algorithm ensure the classification task. As the results showed that deep neural networks are more accurate than random forests. Using the GUI for training and testing classifiers facilitates the operation for several faults diagnosis without a programming background. The obtained results demonstrate the good performance of the suggested approach for bearing fault diagnosis.

REFERENCES

[1] A. Heng, S. Zhang, A. C. C. Tan, and J. Mathew, "Rotating machinery prognostics: State of the art, challenges, and opportunities," Mech. Syst. Signal Process., vol. 23, no. 3, pp. 724739, 2009.

[2] Bonnett, A.H., Yung, C.: Increased efficiency versus increased reliability. IEEE Ind. Appl. Mag. 14(1), 29–36 (2008).

[3] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study," Mech. Syst. Signal Process., vols. 6465, pp. 100131, Dec. 2015.

[4] Sharma, Aditya, et al. “Feature Extraction and Fault Severity Classification in Ball Bearings.” Journal of Vibration and Control, vol. 22, no. 1, Jan. 2016, pp. 176–192, doi:10.1177/1077546314528021.

[5] T. Thelaidjia, S. Chenikher, and A. Moussaoui, "An Effective Approach for Severity Fault Diagnosis of Rolling Bearings," 2019 1st International Conference on Sustainable Renewable Energy Systems and Applications (ICSRESA), Tebessa, Algeria, 2019, pp. 1-6, DOI: 10.1109/ICSRESA49121.2019.9182688.

[6] Wu, D., Wang, H., Liu, H., He, T., Xie, T.: Health monitoring on the spacecraft bearings in high-speed rotating systems by using the clustering fusion of normal acoustic parameters. Appl. Sci. 9(16), 3246 (2019).

[7] Samanta, B., Al-Balushi, K.R.: Artificial neural network-based fault diagnostics of rolling element bearings using time-domain features. Mech. Syst. Signal Process. 17(2), 317–328 (2003).

[8] Kankar, P.K., Sharma, S.C., Harsha, S.P.: Fault diagnosis of ball bearings using machine learning methods. Expert Syst. 1876–1886 (2011).

[9] Marichal, G., Artes, M., Garcia-Prada, J.: An intelligent system for faulty-bearing detection based on vibration spectra. J. Vib. Control 17(6), 931–942 (2011).

[10] Khadersab, A., Shivakumar, S.: Vibration analysis techniques for rotating machinery and its vibration analysis effect techniques for rotating on bearing faults machinery and its effect on bearing faults costing models for capacity optimization in industry 4. 0: trade-off between used capacity and operational efficiency. Procedia Manuf. 20, 247–252 (2018).

[11] Nikolaou, N.G., Antoniadis, I.A.: Rolling element bearing fault diagnosis using wavelet packets. NDT& E Int. 35, 197–205 (2002).

[12] Liu, T.I., Singonahalli, J.H., Iyer, N.R.: Detection of roller bearing defects using expert system and fuzzy logic. Mech. Syst. Signal Process. 10, 595–614 (1996).

[13] Jenkins, C.D.: Bearing fault detection and wear estimation using machine learning. Los Alamos National Lab (LANL), Los Alamos (2019).

[14] Alguindigue, I.E., Loskiewicz-Buczak, A., Uhrig, R.E.: Monitoring and diagnosis of rolling element bearings using artificial neural networks. IEEE Trans. Ind. Electron. 40(2), 209–217 (1993).

[15] Jack, L.B., Nandi, A.K.: Genetic algorithms for feature selection in machine condition monitoring with vibration signals. IEEE Proc. Vis. Image Signal Process. 147(3), 205–212 (2000).

[16] Sawaqed, L.S., Alrayes, A.M. Bearing fault diagnostic using machine learning algorithms. ProgArtifIntell 9, 341–350 (2020). https://doi.org/10.1007/s13748-020-00217-z

[17] W. Li, A. Monti, and F. Ponci, "Fault Detection and Classification in Medium Voltage DC Shipboard Power Systems with Wavelets and Artificial neural networks," in IEEE Transactions on Instrumentation and Measurement, vol. 63, no. 11, pp. 2651-2665, Nov. 2014, DOI: 10.1109/TIM.2014.2313035.

[18] Gs, Vijay &Pai, P. &Sriram, N. S.. (2010). Artificial Neural Network based Condition Monitoring of Rolling Element Bearing.

[19] D. E. Rumelhart and J. L. McClelland, Parallel Distributed Processing. Cambridge, MA, USA: MIT, 1986.

[20] R. P. Lippmann, “An introduction to computing with neural nets,” IEEE ASSP Mag., vol. 4, no. 2, pp. 4–22, Apr. 1987.

[21] Building A Deep Learning Model using Keras | by EijazAllibhai | Towards Data Science.

[22] "IEEE Standard for Transitions, Pulses, and Related Waveforms," in IEEE Std 181-2011 (Revision of IEEE Std 181-2003), vol., no., pp.1-71, 6 Sept. 2011, DOI: 10.1109/IEEESTD.2011.6016198.

[23] MathWorks R2021, Copyright 1993-2004 The MathWorks, Inc. Available: https://www.mathworks.com/help/predmaint/ug/signal-features.html

[24] Dr. ABHISHEK SHARMA, A Simple Analogy to Explain Decision Tree vs. Random Forest, May 12, 2020, Accessed on: December 28, 2020 [Online]. Available: https://www.analyticsvidhya.com/blog/2020/05/decision-tree-vs-random-forest-algorithm

[25] Case Western Reserve University Bearing Data Center Website http://www.eecs.case.edu/laboratory/bearing/download.htm